第26卷 第5期
2009年10月

工 程 数 学 学 报
CHINESE JOURNAL OF ENGINEERING MATHEMATICS

Vol. 26 No. 5
Oct. 2009

文章编号:1005-3085(2009)05-0827-09

Markov 动态决策过程在耐用品购买中的应用*

贾俊秀

(西安电子科技大学理学院，经济管理学院，西安 710071)

摘 要: 通过建立一个离散型Markov决策模型，并在模型中引入感知价值和感知价值系数，本文研究了消费型耐用品购买的多阶段最优策略。当每个产品阶段中有两代不同产品时，消费者最优购买策略存在；在一定的条件下，从某个状态开始最优策略为“购买”。最后以一个数值实例验证了这些结论，并说明给定相关数据后可获得购买耐用品的最优决策，这一决策与消费者的偏好和卖方的定价存在密切关系。

关键词: Markov 决策过程；耐用消费品；最优购买决策；感知价值

分类号: **AMS**(2000) 60J20　　**中图分类号**: O211.62; F224　　**文献标识码**: A

1 引言

不管是消费型耐用品还是工业型耐用品，它们在商业周期的产生和发展中起着重要作用，在国民生产总值中占据较大比例。因此，很多学者都非常关注耐用品理论的研究和发展，如文献 [1-3]。文献 [4-6] 从生产商的角度研究了耐用品的定价问题。文献 [7] 也从卖方的角度研究了商品动态价格的数学模型，并给出动态价格最优的条件。本文则从消费者的立场研究耐用品的最优购买决策，相关文献概述如下。

Bass研究了新消费品的首次购买决策理论，从创新和模仿行为的角度提供了一种理性行为，并假设消费者的首次购买决策与先前购买者的数量有关。Coase认为如果科斯猜想成立，那么只有当耐用品价格等于其边际费用的那个时刻才会有交易产生。尽管现实生活中没有这种极端情形，但随着产品生命周期的继续产品价格会降低，对于耐用品来说也是很普遍的事情[8]。本文也作了这样的假设，认为随着时间的推移，生产商会降低旧产品的价格。这是一个共同知识，消费者也知晓。而且，一般情况下今天购买某种耐用品的顾客不会明天再去买同一型号的产品，因此同一个生产商在不同阶段上市的产品为替代品，而非互补品。同时本文在模型中引入了消费者感知价值和感知价值系数。对此，Banerjee和Bandyopadhyay[9]、Wathieu[10] 利用感知价值的概念分析了消费者的行为。文献 [11] 则从网络消费者的角度，结合感知价值对消费者的决策行为，运用结构方程模型方法进行了定量研究。研究结果表明，感知价值与购买意愿显著正相关。

对于一些消费者来说，尽管消费型耐用品的贬值率很高，但购买新的替代品可以立即享受多功能、高性能和质量的产品，只是需要支付较高的价格。生产商通常会迎合消费者的这一心理，尽力引入新产品，赚取更多利润，如中国的手机市场，就是这种情形。而另外一些消费者可能会以低价购买现有旧款产品，或寻找性价比高的产品。因此，消费者的决策问题是什么时间购买产品自己会有更高的收益或效用。程兰芳[12] 根据价格的实物期权性研究了这一问题，

收稿日期: 2008-02-27.　**作者简介**: 贾俊秀 (1974年8月生)，女，博士，讲师. 研究方向：消费者行为与供应链管理.
***基金项目**: 国家自然科学基金 (70701028).

给出购买产品关键时刻的分析规则。Dhebar[13] 对耐用品垄断者相继出售不同版本产品 (两代产品) 的问题建立了两阶段模型，假设没有二手市场。在他的模型中，生产商在第一阶段销售第一代产品，在第二个阶段销售第二代产品；消费者是理性的，目的是最大化收益减去价格的差额。而更一般的情况为消费者在适当的时间点上对多阶段替代产品进行选择，而消费理论的恰当基础就是研究现实消费和未来消费之间交替关系的跨时选择理论[14]。

本文则给出两代产品、多阶段的耐用品购买决策模型。特别适用于那些高价的耐用消费品，如摄像机、笔记本电脑和 DVs 等产品。对于某个消费者来说，他购买产品后将永远退出市场，没购买产品则进入下一个阶段进行重新决策。为此，我们建立了一个离散性 Markov 决策过程模型描述消费者的购买过程。消费者对不同代产品的保留价作为系统的状态；报酬函数由两部分组成，一部分为产品的物理价值，另一部分为消费者的感知价值；目标函数为最大化消费者在耐用品购买过程中的总体收益；目的是要寻找消费者的最优购买策略。实例分析了具有两代产品多个阶段的情况，说明消费者的最优购买策略除了与厂家定价、保留价格和感知价值有关外，与消费者的偏好也存在密切关系，即消费者的类型在决策中起很重要的作用。本文为首次应用 Markov 决策过程模型对耐用品消费者的购买决策进行建模分析，希望能给消费者和生产企业以决策支持。

2 模型建立

首先给出一个合理的假设：消费者最多购买一件耐用产品，这是由于产品的耐用性和高价性。另一个假设为消费者不进入二手市场[4]，这是为了集中精力分析相继上市产品对消费者决策的影响。同时，模型中的状态转移概率是固定的，文中所用其它记号解释如下。

n：文中称其为产品阶段，可以理解为新旧产品出现的间隔时间。如果每个产品阶段中有多代产品，比如 m 代不同的产品，消费者在每个产品阶段必须进行 m 次购买决策，即消费者共需要决策 mn 次，我们将其称为购买阶段，可知总共有 mn 个购买阶段；

s：消费者对某代产品给出的保留价。产品不同，保留价也不同。$S=\{s_1,s_2,\cdots,s_{nm}\}$ 为有限的状态空间，其中 $s_{nm}=0$ 表示消费者未购买任何产品而离开市场；

$T(s'|s)$：状态转移概率，s'，$s\in S$；

$a(s)$：消费者在状态 s 处选择的行动，从集合 $\{B,W\}$ 中取值。B (Buy it) 表示购买产品并退出市场，W (Wait) 表示继续留在市场中等待购买合适自己的产品；

$r(s,B)$：消费者在状态 s 处购买产品时获得的当阶段总折扣期望回报。实际上，$r(s,B)$ 同时是产品价格、性能、质量、消费者类型(风险中性、偏好和厌恶)和消费者虚荣程度等的函数。为方便书写表示为 $r(s,B)$，另一个原因是在文中这些因素多数被假设为常数。$r(s,W)$ 为消费者没有购买产品时，在当前阶段中累计的期望回报；

$v_{p.v.}(s,B)$：消费者购买产品获得的实际物质收益；$v_{p.g.}(s,B)$ 为感知收益，$v_{p.l.}(s,W)$ 为感知损失；α 为感知系数；

λ：折扣率，$0\le\lambda\le 1$；

$V(s)$：消费者在状态 s 处获得的最大总折扣期望回报。求解如下方程可得到问题的最优解

$$V(s)=\max\Big\{r(s,B),\, r(s,W)+\lambda\sum_{s'\in S}T(s'|s)V(s')\Big\},\quad s\in S, \tag{1}$$

其中

$$r(s,B)=(1-\alpha)v_{p.v.}(s,B)+\alpha v_{p.g.}(s,B),\quad r(s,W)=s^b+v_{p.l.}(s,W). \tag{2}$$

在当前社会下，对于耐用品来说，特别是那些耐用电器产品，如手机、笔记本电脑和汽车等，消费者购买某款产品的动机有两部分组成，主要的一部分是因为耐用品本身有实用价值，另一部分则来自于消费者的虚荣心，如攀比、赶时髦等心理。也就是说消费者购买产品可以得到物质上的和精神上的双重收益。物质上的用 $v_{p.v.}(s,B)$，s^b 来反映，精神上的用 $v_{p.g.}(s,B)$，$v_{p.l.}(s,W)$ 来反映。为了体现消费者的上述双重收益，我们引入参数“感知系数 α”，用来反映感知收益在消费者总收益中所占的比例，因此 α 能反应消费者的类型。本文确定此参数的方法与文献 [9,10] 不同，可以很好的体现消费者的主观意志在购买决策中的重要性。反之，如果消费者没有购买产品，他会获得感知损失。除此之外，因为没有购买产品而节省下来的资金算作不购买产品的实际收益，用厂家标价 s^b 来体现。其中，$v_{p.g.}$，$v_{p.l.}$ 可通过 Von Neumann 和 Morgenstern 的心理测验法获得，也可应用本文给出的方法确定；需要解释的是：感知收益和感知损失相对于产品的实际功能价值来说小得多，但它们会影响消费者的购买决策。感知系数 α 可通过 Savage (1954) 提出的主观概率法[15] 来确定，α 随着购买阶段的变化而变化。

3 多阶段最优购买决策

我们认为消费者对某代产品的保留价是固定的，但产品不同保留价不同。除此之外，还有如下重要假设：

假设 1 耐用品型号越新，消费者给出的保留价越高。

假设 2 产品性价比和质量降低时，消费者购买折扣收益 $r(s,B)$ 不会增加。

假设 3 消费者不购买产品时的 $v_{p.l.}(s,W)$ 不是 s 的增函数。

这些假设都是基本符合现实世界的实际情况，很容易被理解，所以不给出进一步的说明。现在重点分析 $m=2$ 的情形，也就是说在每个产品阶段中有两代不同产品可供消费者决策。产品推出的规则为：当新一代产品上市时，它的上上一代产品将退出市场，如图 1 所示。

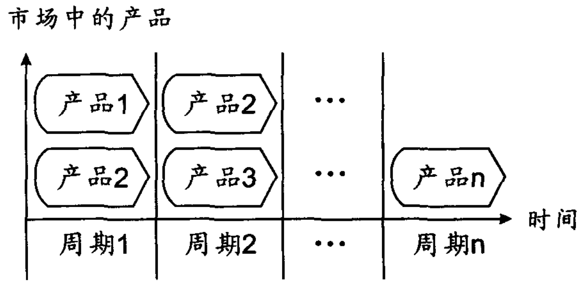

图 1: 产品推出过程—每个阶段中有两代产品

在图 1 中，产品 $k(k=2,3,\cdots,n)$ 在阶段 $k-1$ 开始上市，在下一个阶段末退出市场；第一个阶段中有两代产品，产品 1 为早期遗留产品，在下一个阶段退出市场；最后一个阶段中仅有一个产品，来自上一个阶段出现的新产品。又因为不同阶段中同一种型号产品报价不同，即对于同一型号产品，消费者可能在不同购买期进行选择，因此包括吸收态共有 $2n$ 个状态；很容易写出公式 (3) 所示的状态转移矩阵，其中的 s_{2n} 为吸收态。比如，在第一个产品阶段中消费者需要决策两次，首先决定是否购买产品 1，如果不买的话，可以转移到后面的各个购买期。如第二购买期的产品 2、第三购买期的产品 2 等等，这两个产品虽然型号相同但可能价格

不同。又因为大多被调查的消费者很难说清楚从某一确定状态转向其它状态的概率，所以此处假设这个概率是相同的，且为平均数，如公式(3)所示。

$$T = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_{2n-1} \\ s_{2n} \end{array} \begin{array}{c} \begin{array}{cccccc} s_1 & s_2 & s_3 & \cdots & s_{2n-1} & s_{2n} \end{array} \\ \begin{pmatrix} \frac{1}{2n} & \frac{1}{2n} & \frac{1}{2n} & \cdots & \frac{1}{2n} & \frac{1}{2n} \\ \frac{1}{2n} & \frac{1}{2n} & \frac{1}{2n} & \cdots & \frac{1}{2n} & \frac{1}{2n} \\ 0 & \frac{1}{2n-1} & \frac{1}{2n-1} & \cdots & \frac{1}{2n-1} & \frac{1}{2n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \end{array}. \tag{3}$$

现在，要求解的问题是

$$\begin{aligned} V(s) \ = \ \max\Big\{&(1-\alpha)v_{p.v.}(s,B)+\alpha v_{p.g.}(s,B), s^b \\ &+v_{p.l.}(s,W)+\lambda\sum_{s'\in S}T(s'|s)V(s')\Big\}, \quad s\in S=\{s_1,s_2,\cdots,s_{2n}\}. \end{aligned} \tag{4}$$

假设消费者购买产品获得的物理价值可由实际的产品出售价格 s^p 来衡量，因为这个价格是消费者认同的产品价值。其中，s^p 表示状态 s 处购买产品的实际价格。而在某个阶段中某个状态 s 下购买产品的感知价值和未购买到产品的损失价值由如下表达式确定

$$v_{p.g.}(s,B)=\beta\ln s, \quad v_{p.l.}(s,W)=\gamma\ln s, \quad s>1. \tag{5}$$

其中 $s>1$ 可以通过变化其单位得以保证。此时，优化问题可以写成

$$V(s_k)=\begin{cases} \max\Big\{(1-\alpha_k)s_k^p+\alpha_k\beta\ln s_k,\ s_k^b+\gamma\ln s_k \\ \qquad +\lambda\sum\limits_{s'\in S}\frac{1}{2n+2-k}V(s')\Big\}, \quad k=3,4,\cdots,2n-1, \\ \max\Big\{(1-\alpha_k)s_k^p+\alpha_k\beta\ln s_k,\ s_k^b+\gamma\ln s_k+\lambda\sum\limits_{s'\in S}\frac{1}{2n}V(s')\Big\}, \quad k=1,2. \end{cases} \tag{6}$$

边界条件为 $V(s_{2n})=0$。其中 $\beta>0$ 说明消费者购买产品后的感知价值会随着保留价的增加而增加，但增加的幅度是由对数函数来决定的。同理 $\gamma<0$ 表示消费者在给定阶段内没有购买到产品的整体价值会随着保留价的增加而降低，其幅度是由对数函数来决定的。对于某个消费者来说，β 和 γ 的值相对固定。

由假设1知，消费者对相继出现的新产品的保留价是越来越高的，总会有一个时刻会使消费者等待价值的降低会高于购买产品价值的增加。所以，我们有如下定理。

定理1 如果消费者从一个购买期转移到其它购买期的转移概率相同，且感知价值和损失价值都是保留价的对数函数，在每个产品阶段有两代产品的情形下，若如下条件成立，

$$\begin{aligned} &1-\frac{(1-\alpha_{k+1})s_{k+1}^p+\alpha_{k+1}\beta\ln s_{k+1}+s_k^b-s_{k+1}^b}{(1-\alpha_k)s_k^p+\alpha_k\beta\ln s_k} \\ &\le \lambda\Big[\frac{1}{2n+2-k}-\frac{1}{2n+1-k}\Big], \quad s_k=s_l,\ s_{l+1},\ \cdots,\ s_{2n-2}. \end{aligned} \tag{7}$$

则存在某个状态 s_l，使

$$a^*(s_1)=a^*(s_2)=\cdots=a^*(s_{l-1})=W,$$

$$a^*(s_l)=a^*(s_{l+1})=\cdots=a^*(s_{2n-1})=B.$$

简单的证明思路为：如果能证明对于任一 l，当 $a^*(s_l)=B$，则 $a^*(s_{l+1})=B$，则定理可以得证，所以应用反正法假设存在某个 l，当 $a^*(s_l)=B$ 时，$a^*(s_{l+1})=W$。如果定理中条件成立，则推出

$$V(s_l)<(1-\alpha_k)s_l^p+\alpha_k\beta\ln s_{l+1},$$

即 $V(s_l)<r(s_l,B)$ 这一结论与假设矛盾。从而可得到在 l 处的最优决策为 B 时，$l+1$ 处的最优策略也应该为 B，证明如下。

证明 用反证法。

假设存在某个 l，当 $a^*(s_l)=B$ 时，$a^*(s_{l+1})=W$。当 $l\geq 3$ 时，则有

$$\begin{aligned}(1-\alpha_l)s_l^p+\alpha_l\beta\ln s_l &> \gamma\ln s_l+\lambda\sum_{s'\in S}\frac{1}{2n+2-l}V(s')+s_l^b\\ &=\gamma\ln s_l+\lambda\frac{1}{2n+2-l}\sum_{s'\in S}V(s')+s_l^b,\end{aligned}\tag{8}$$

$$\begin{aligned}(1-\alpha_{l+1})s_{l+1}^p+\alpha_{l+1}\beta\ln s_{l+1} &< \gamma\ln s_{l+1}+\lambda\sum_{s'\in S}\frac{1}{2n+2-l-1}V(s')+s_{l+1}^b\\ &=\gamma\ln s_l+\lambda\frac{1}{2n+2-l-1}\sum_{s'\in S}V(s')+s_{l+1}^b.\end{aligned}\tag{9}$$

所以

$$\begin{aligned}&(1-\alpha_l)s_l^p+\alpha_l\beta\ln s_l-(1-\alpha_{l+1})s_{l+1}^p-\alpha_{l+1}\beta\ln s_{l+1}\\ >&\ \gamma\ln s_l-\gamma\ln s_{l+1}+\lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)\sum_{s'\in S}V(s')+s_l^b-s_{l+1}^b\\ \geq&\ \lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)\sum_{s'\in S}V(s')+s_l^b-s_{l+1}^b\\ =&\ \lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)\sum_{s'\geq s_l}V(s')+\lambda\frac{1}{2n+2-l}V(s_{l-1})+s_l^b-s_{l+1}^b\\ \geq&\ \lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)\sum_{s'\geq s_l}V(s')+s_l^b-s_{l+1}^b\\ \geq&\ \lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)V(s_l)+s_l^b-s_{l+1}^b,\end{aligned}\tag{10}$$

即

$$\begin{aligned}&(1-\alpha_l)s_l^p+\alpha_l\beta\ln s_l-(1-\alpha_{l+1})s_{l+1}^p-\alpha_{l+1}\beta\ln s_{l+1}-(s_l^b-s_{l+1}^b)\\ >&\ \lambda\left(\frac{1}{2n+2-l}-\frac{1}{2n+2-l-1}\right)V(s_l),\end{aligned}$$

其中从公式 (10) 第一步到第二步是因为 $\gamma < 0$、$\ln s_l$ 单调不减，所以 $\gamma\ln s_l - \gamma\ln s_{l+1} \geq 0$；从公式 (10) 第三步到第四步是因为每个购买期中的最大收益不会为负值。

如果

$$\frac{(1-\alpha_l)s_l^p + \alpha_l\beta\ln s_l - (1-\alpha_{l+1})s_{l+1}^p - \alpha_{l+1}\beta\ln s_{l+1} - (s_l^b - s_{l+1}^b)}{(1-\alpha_l)s_l^p + \alpha_l\beta\ln s_l}$$

$$\leq \lambda\left[\frac{1}{2n+2-l} - \frac{1}{2n+1-l}\right],$$

即

$$(1-\alpha_l)s_l^p + \alpha_l\beta\ln s_l - (1-\alpha_{l+1})s_{l+1}^p - \alpha_{l+1}\beta\ln s_{l+1} - (s_l^b - s_{l+1}^b)$$

$$\leq \lambda\left[\frac{1}{2n+2-l} - \frac{1}{2n+1-l}\right]r(s_l, B),$$

则有 $V(s_l) < r(s_l, B)$，这与假设矛盾。假设错误，定理 1 成立。当 $l = 1, 2$ 时可同理证明。

4 模型应用

本节应用某个消费者准备购买某种品牌手机的数据，讨论第 3 节中模型的求解情况。假设以一个月为一代产品推出的阶段，即产品阶段，表 1 为连续三个月的调查数据。也就是说，消费者每个月都需要作出 2 次是否购买产品的决策，三个月后退出市场。其中，Q_1 表示对消费者 1 的调查数据，Q_2 表示对消费者 2 的调查数据，当然可以对任意一个消费者的数据进行分析。V_1, V_2, V_3 表示出现三代产品，在第一个阶段中有 V_1, V_2 两代产品存在；在第二个阶段中有 V_2, V_3 两代产品存在，V_1 已经退出市场；假设第三个月中只有产品 V_3 存在。标价仅是一个参考价格，实际购买价会上下波动，因此保留价低于这个标价并不意味着消费者不购买此产品。表中的感知系数反应消费者的类型，这一点可从很多中国手机消费者的行为看出，他们除了注重手机的实际功能外，更多地将手机看作一种装饰品，更新换代频率令发达国家的人都咂舌。

表 1: 对不同耐用品消费者的调查数据 (价格单位：元)

	决策阶段	第一个月		第二个月		第三个月
	产品类型	V_1	V_2	V_2	V_3	V_3
	产品标价 s_k^b	2180	3070	3030	3140	3100
Q_1	消费者保留价格 s_k	2028	3020	3020	3030	3030
	感知系数 α	0.34	0.15	0.08	0.08	0.07
Q_2	消费者保留价格 s_k	1900	2300	2000	2450	2100
	感知系数 α	0.16	0.10	0.12	0.18	0.14

其中感知系数的取值是根据 Savage 提出的主观概率法来确定的。首先，设计问题和与问题相关的两个事态体，包括消费者所认同的、有关感知收益在消费者总收益中所占的比例 (感知系数) 的原事态体和用来逐步度量消费者对问题的真实回答的参考事态体。然后由消费者在两个

事态体间选择辨优，根据消费者的回答，逐步修改参考事态体，直到得到与感知价值等同的相信程度，即认为这样得到的主观概率就是我们的感知价值。

4.1 分析求解

如下应用对消费者1的调查数据，从状态、行动、转移概率、报酬函数和目标函数几个方面分析消费者进行耐用品购买的Markov决策过程。

- 状态：由前面对阶段的解释和对表1的分析可知，此例中有6个状态，其中 s_6 是一个吸收态；
- 行动：消费者可从 $\{B, W\}$ 中选择行动；
- 转移矩阵：如公式 (11) 所示；
- 收益函数：如第3节中描述的那样，可同时体现消费者的实际收益和感知收益，s_k，s_k^b，α 的具体数值见表1；$s_k^p, k=1,\cdots,5$ 的值分别为2040、3056、3030、3080、3035元；假设此消费者的 $\beta=86$，$\gamma=316$；
- 目标函数：最大化消费者总的折扣期望价值。在本例中，由于每个阶段的持续时间较短，所以假设折扣系数 $\lambda=1$。

$$
T = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{array}
\begin{array}{c}
\begin{array}{cccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{array} \\
\begin{pmatrix}
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
0 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\
0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\
0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}. \tag{11}
$$

4.2 计算过程与结果

当 $n=6$ 时，$V(s_6)=0$。

当 $n=5$ 时有

$$
\begin{aligned}
V(s_5) &= \max\Big\{(1-0.07)3035+0.07\times 86\times\ln 3030, \\
&\qquad 3100-316\ln 3030+\frac{1}{3}(V(s_4)+V(s_5)+V(s_6))\Big\} \\
&= \max\Big\{2870.81, 566.84+\frac{1}{3}(V(s_4)+V(s_5))\Big\} \\
&= \begin{cases} 2870.81, & a^*(s_5)=B, \\ 566.84+\frac{1}{3}(V(s_4)+V(s_5)), & a^*(s_5)=W \text{ 且 } V(s_4)>3956.974. \end{cases}
\end{aligned} \tag{12}
$$

也就是说，当 $a^*(s_5)=W$ 时，有

$$
V(s_5)=566.84+\frac{1}{3}(V(s_4)+V(s_5))>2870.81,
$$

可推出在这种情况下，则 $V(s_4)$ 应该大于3956.974。

当$n=4$时，同理可求得

$$V(s_4) = \max\{2888.75, 606.84 + \frac{1}{4}(V(s_3) + V(s_4) + V(s_5))\}$$

$$= \begin{cases} 2888.75, & a^*(s_4) = B, \\ 606.84 + \frac{1}{4}(V(s_3) + V(s_4) + V(s_5)), & a^*(s_4) = W \text{ 且 } V(s_3) > 3368.074. \end{cases}$$

由上述数据可知，如果$a^*(s_4)=B$，则一定有$a^*(s_5)=B$，因为此时

$$V(s_4) = 2888.75 < 3956.974.$$

通过进一步计算$V(s_3), V(s_2), V(s_1)$的值，可确定$P=\{W,B,B,B,B\}$为此消费者的最优策略。此结论也可通过定理1进行验证。这些结论都是基于消费者给出的实际数据而得出，因此本文提供的方法可充分体现消费者决策的主观能动性，满足消费者主客观上对物品的需求。我们计算了部分其它问卷的数据，得到的最优策略当然都不太相同。但这些结果都反应了一个事实：最优策略除了与商家的定价、消费者给出的保留价、感知价值有关外，与体现消费者类型的α也密切相关。如果企业能掌握了大量的消费者类型和购买决策的信息，那么就能更加准确地预测需求，从而准确制定生产、销售等决策，故此项研究不仅有助于消费者的决策，同时可帮助企业进行决策。

5 结论

本文用Markov决策过程模型描述了消费者购买耐用品的动态过程。为了更好地体现消费者决策的主观能动性，对传统的Markov决策过程模型作了适当改动。在模型中引入了感知系数这个新概念、感知收益和感知损失。在多阶段情况下，重点给出了每个阶段中有两代产品时，消费者的最优购买策略。在满足给定条件的前提下，消费者的最优策略为：存在一个最优购买时刻s_l，从这个时刻开始，只要消费者还停留在市场中，他的最优决策是“购买”。文章还用一个实例详细给出求解最优策略的过程。需要进一步探讨的问题是，如果每个阶段中有多于两代的产品出现，如何求解消费者的最优购买策略；还有物理价值和感知价值的不同度量方法的比较研究等问题。

参考文献：

[1] Bass F M. A new product growth for model consumer durables[J]. Management Science, 2004, 50(12): 1825-1832

[2] Coase R H. Durability and monopoly[J]. Journal of Law and Economics, 1972, 15: 143-149

[3] Huang S, Yang Y, Anderson K. A theory of finitely durable goods monopoly with used-goods market and transaction costs[J]. Management Science, 2001, 47(11): 1515-1532

[4] Kornish L J. Pricing for a durable goods monopolist under rapid sequential innovation[J]. Management Science, 2001, 47(11): 1552-1561

[5] Ray S, Boyaci T, Aras N. Optimal prices and trade-in rebates for durable, remanufacturable products[J]. Manufacturing and Service Operations Management, 2005, 7: 208-228

[6] Chien H-K, Chu C Y. Sale or lease? Durable-goods monopoly with network effects[J]. Marketing Science, 2008, 27(6): 1012-1019

[7] 化存才. 价格波动时商品动态价格的数学模型、物价稳定性与宏观调控分析[J]. 工程数学学报，2007, 24(3): 447-450

[8] Balachander S, Srinivasan K. Modifying customer expectations of price decreases for a durable product[J]. Management Science, 1998, 44(6): 776-786
[9] Banerjee B, Bandyopadhyay S. Advertising competition under consumer inertia[J]. Marketing Science, 2003, 22(1): 131-144
[10] Wathieu L. Consumer habituation[J]. Management Science, 2004, 50(5): 587-596
[11] 王崇，李一军，叶强. 互联网环境下基于消费者感知价值的购买决策研究[J]. 预测，2007, 26(3): 21-26
[12] 程兰芳. 居民家庭耐用消费品的消费时机选择：实物期权的方法[J]. 运筹与管理，2004, 13(1): 126-129
[13] Dhebar A. Durable-goods monopolists, rational consumers and improving products[J]. Marketing Science, 1994, 13(1): 100-120
[14] Deaton A(著). 胡景北，鲁昌(译). 理解消费[M]. 上海：上海财经大学出版社，2003
[15] 李华，胡奇英. 预测与决策[M]. 西安：西安电子科技大学出版社，2005

Markov Dynamic Decision Model for Purchase Decision of Durable Goods

JIA Jun-xiu

(School of Science and School of Economics and Management, Xidian University, Xi'an 710071)

Abstract: A discrete time Markov decision process model was built, introduced in the model are the concepts of perceived value and perceived coefficient. The optimal policies to purchase consumptive durable goods were studied under multi-stages. When there are two generations in each product stage, the consumer's optimal purchasing policy exists. It is proved that exists some state, from which the consumer's optimal policy is "to buy the product" under some conditions. A numerical example show the reliability of the method. Given related data, the optimal purchase policies can be obtained, which have closed relation with the consumer's preference and the pricing of the bargainor.

Keywords: Markov decision process; durable goods; optimal purchase decision; perceived value